Supporting Information:

Sequence, Structure, and Function of Peptide Self-assembled Monolayers

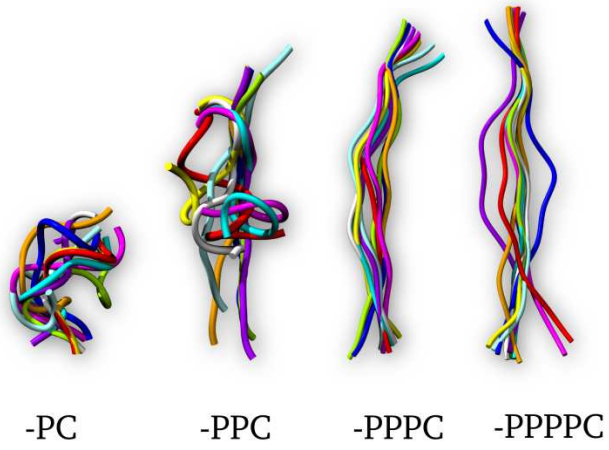
Ann K. Nowinski, Fang Sun, Andrew D. White, Andrew J. Keefe, and Shaoyi Jiang*

Department of Chemical Engineering, University of Washington, Seattle, Washington 98195, United States

1. Additional Experimental Data

1a. Effect of peptide linker length on protein adsorption.

a)



b)

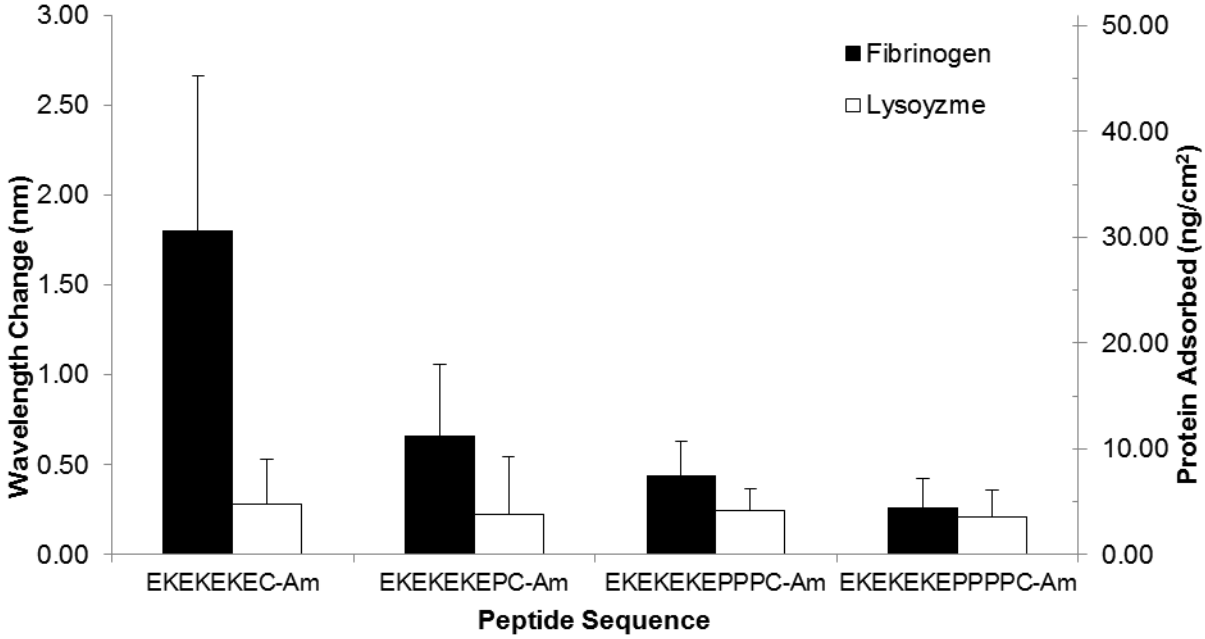


Figure S1. Effect of peptide linker length on protein adsorption. (a) Ten lowest energy (score) conformations from 24 different starting structures as predicted using the Rosetta structure prediction package. They are aligned and oriented with the cysteine (C-terminus) at the bottom. The important feature to note is the increasing extended conformations as proline is added, due to the change from random coil to polyproline helix. The proline also makes the conformations more rigid, which should reduce the entropic penalty of peptide adsorption onto a gold surface. Sequences from left to right are EKEKEKEPC-Am, EKEKEKEPPC-Am, EKEKEKEPPPC-Am, and EKEKEKEPPPPC-Am. (b) Adsorption of protein on peptide SAMs composed of proline linker series EKEKEKEC-Am, EKEKEKEPC-Am, EKEKEKEPPPC-Am and EKEKEKEPPPPC-Am determined from a surface plasmon resonance sensor (SPR) in the unit of wavelength shift (nm) or converted surface concentration (ng/cm^2). (Black – fibrinogen, White – lysozyme). Each data point represents an average value ± standard deviation (SD) from at least three independent measurements.

1b. XPS of peptide SAMs.

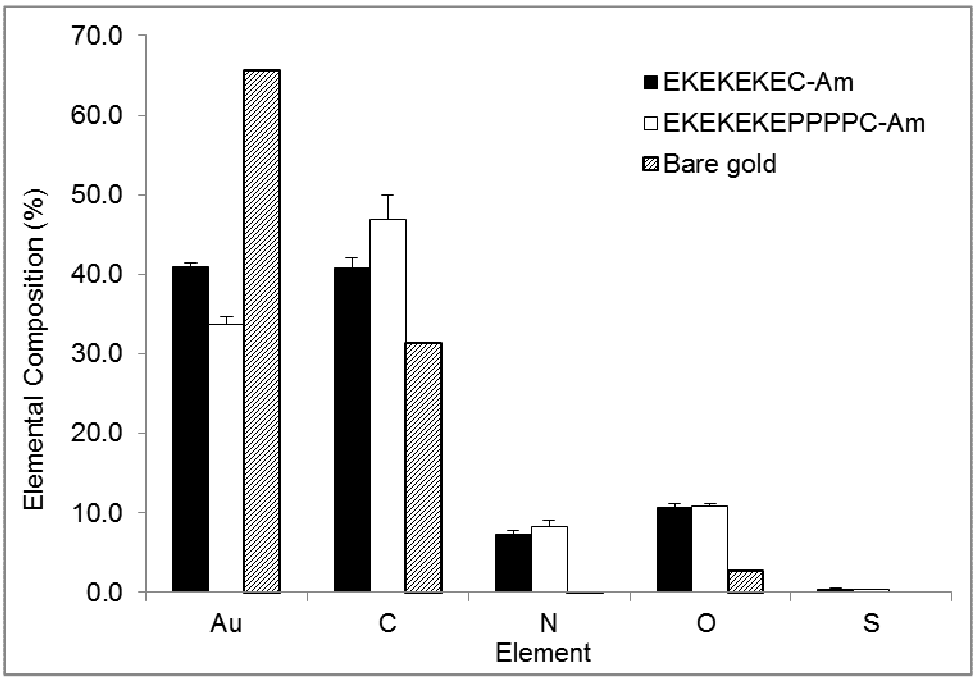


Figure S2. XPS surface compositions. Percent elemental surface compositions of Au, C, N, O, and S for EKEKEKEC-Am (Black) and EKEKEKEPPPPC-Am (White) peptide SAMs and bare gold control (Stripes).

	EKEKEKEC-Am		**EKEKEKEPPPPC-Am**	
	Theoretical	Actual	Theoretical	Actual
C	58.6	68.8	62.2	70.6
N	17.1	12.5	16.3	12.5
O	22.9	18.1	20.4	16.4
S	1.4	0.6	1.0	0.5

Table S1. XPS theoretical vs actual elemental percent compositions for C, N, O, and S. Calculated theoretical percentages for elements carbon, nitrogen, oxygen, and sulfur for peptide SAMs versus actual experimental percentages obtained.

1c. Control cell adhesion experiment with scrambled RDGEKEKEKEPPPPC-Am peptide.

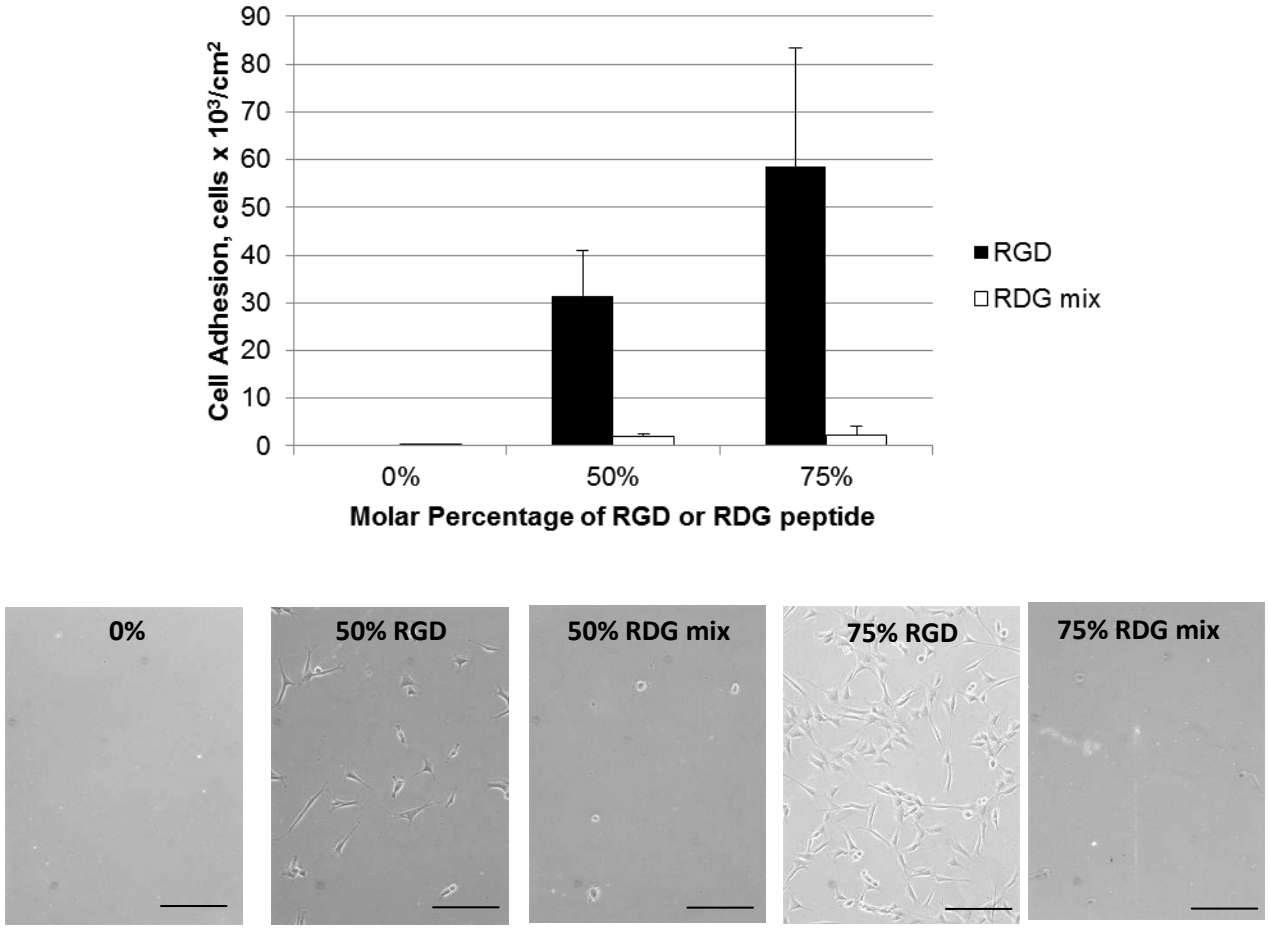


Figure S3. Control experiment showing specific cell adhesion is achieved. Cell adhesion results for peptide SAMs composed of mixtures of RGDEKEKEKEPPPPC-Am and the scrambled sequence RDGEKEKEKEPPPPC-Am with EKEKEKEPPPPC-Am. The molar percentage of the RGD peptide and scrambled RDG peptide tested was 0%, 50%, and 75%. Scale bars represent 100 μm. The same cell seeding procedure of peptide SAMs was followed for the control experiment as listed in the paper, with the exception that the cell plating concentration was 3.3 x 10^4 cells/mL in 3 mL Dulbecco's Modified Eagle Medium (DMEM) with 10% (v/v) fetal bovine serum (FBS) and 1% (v/v) penicillin-streptomycin (PS).

2. Molecular Simulation Methods

The temperature distribution for replica exchange is: 300.0, 301.2, 302.5, 303.7, 305.0, 306.2, 307.5, 308.7, 310.0, 311.3, 312.5, 313.8, 315.1, 316.4, 317.7, 319.0, 320.3, 321.6, 323.0,

324.3, 325.6, 326.9, 328.3, 329.6, 331.0, 332.3, 333.7, 335.1, 336.5, 337.8, 339.2, 340.6, 342.0, 343.4, 344.8, 346.2, 347.7, 349.1, 350.5, 352.0, 353.4, 354.9, 356.3, 357.8, 359.2, 360.7, 362.2, 363.7, 365.2, 366.7, 368.2, 369.7, 371.2, 372.7, 374.3, 375.8, 377.3, 378.9, 380.4, 382.0, 383.6, 385.1, 386.7, 388.3, 389.9, 391.5, 393.1, 394.7, 396.3, 398.0, 399.6, 401.2, 402.9, 404.5, 406.2, 407.9, 409.5, 411.2, 412.9, 414.6, 416.3, 418.0, 419.7, 421.5, 423.2, 424.9, 426.7, 428.4, 430.2, 431.9, 433.7, 435.5, 437.3, 439.1, 440.9, 442.7, 444.5, 446.3, 448.2, 450.0.

Single peptides were built using the α-helical Dunbrack rotamers[1]. They were then energy minimized using 10,000 steps of steepest descent. The minimized system was solvated in a box such that the peptide had 1 nm of solvent in all directions. Ions were added to charge equilibrate the system. Annealing was then done for 200 ps with a schedule from 100 K to 500 K in the NPT ensemble. Finally, an equilibration for 100 ps again in the NPT ensemble was conducted. The resulting system was then used for the production simulation described in text.

The “crystallographers’ intuition” codes were used from the DSSP calculations to create the plot in Fig. 3c. There, the mode of the secondary structure is shown visually. The mode corresponds to the free energy minimum when canonically sampling from an NVT ensemble. The lower panel in Fig. 3c was plotted using a kernel density estimator with a bandwidth of 2Å. The choice of a kernel density estimator, as opposed to a histogram, was to allow plotting of the three sequences simultaneously. The horizontal lines are the raw data medians.

The study of addition of proline to the peptides was done using the Rosetta protein structure prediction package[2]. It has had excellent performance in the past with *de novo* structure prediction[2]. The goal of our simulations was not to determine the structure of the peptides, but to instead see if the peptides have a more extended conformation. The peptides were first built to have their ideal rotamers with ideal α-helix rotamers and ideal β-sheet (non-proline residues only), thus giving two starting structures for each sequence. Next, their bonds, angles, and dihedrals were optimized according to the Rosetta scoring function using the “idealize” module of Rosetta. Then, a Monte Carlo optimization of the structures was done for 400,000 steps. The

Monte Carlo moves are backbone dihedral changes. The top 12 structures for the α-helix and β-sheet were chosen for continuation, resulting in 24 structures for each sequence. Next, the "relax" module of Rosetta was used to minimize side-chain clashes by changing χ-rotamers of the sidechains of the top 24 structures. This was followed by another backbone Monte Carlo optimization on each of the 24 structures 300,000 steps and the structures were relaxed. Finally, the 10 structures with the lowest scores among the 24 candidates were aligned and are shown in Fig. S-1b.

1. Dunbrack, R.L. & Cohen, F.E. Bayesian statistical analysis of protein side-chain rotamer preferences. *Protein science□: a publication of the Protein Society* **6**, 1661-81 (1997).
2. Rohl, C. a, Strauss, C.E.M., Misura, K.M.S. & Baker, D. Protein structure prediction using Rosetta. *Methods in enzymology* **383**, 66-93 (2004).